



B-Cell Epitope Mapping from Eight Antigens of *Candida albicans* to Design a Novel Diagnostic Kit: An Immunoinformatics Approach

Farzaneh Vahedi¹ · Younes Ghasemi^{2,3} · Amir Atapour¹ · Kamiar Zomorodian⁴ · Maryam Ranjbar¹ · Ahmad Monabati⁵ · Navid Nezafat^{2,3}  · Amir Savardashtaki^{1,6}

Accepted: 27 April 2022 / Published online: 27 May 2022
© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

Invasive candidiasis is an emerging fungal infection and a leading cause of morbidity in health care facilities. Despite advances in antifungal therapy, increased antifungal drug resistance in *Candida albicans* has enhanced patient fatality. The most common method for *Candida albicans* diagnosing is blood culture, which has low sensitivity. Therefore, there is an urgent need to establish a valid diagnostic method. Our study aimed to use the bioinformatics approach to design a diagnostic kit for detecting *Candida albicans* with high sensitivity and specificity. Eight antigenic proteins of *Candida albicans* (HYR1, HWP1, ECE1, ALS, EAP1, SAP1, BGL2, and MET6) were selected. Next, a construct containing different immunodominant B-cell epitopes was derived from the antigens and connected using a suitable linker. Different properties of the final construct, such as physicochemical properties, were evaluated. Moreover, the designed construct underwent 3D modeling, reverse translation, and codon optimization. The results confirmed that the designed construct could identify *Candida albicans* with high sensitivity and specificity in serum samples of patients with invasive candidiasis. However, experimental studies are needed for final confirmation.

Keywords *Candida albicans* · Epitope mapping · B-cell epitope · Diagnostic kit · Antigens

Introduction

Candida is a significant cause of fungal infections, with *Candida albicans* responsible for 70–90% of candidiasis infections worldwide (Bhakt et al. 2018; Erwig and Gow 2016). As the natural flora of the human body, *Candida albicans* is found on different parts of the body, such as mucosal surfaces, skin, mouth, gastrointestinal tract, and genitals (Antinori et al. 2016; León et al. 2014). The pathogenicity of *Candida albicans* is due to the transformation from single-celled yeast to long multicellular fibers called hyphae. The filamentous form can bind and penetrate the host tissue faster than the single-cell form. Therefore, the production of hyphae indicates the beginning of infection (Nobile et al. 2012; Sachivkina et al. 2020; Tsui et al. 2016). Today, the prevalence of invasive candidiasis as an opportunistic clinical infection is increasing for different reasons (Chen et al. 2020; Lewis and Williams 2017; Thomson et al. 2015). Due to pregnancy and the use of contraceptive estrogens, women are at high risk of fungal infection. *Candida albicans* have been isolated from vaginal biopsies of 20–50% of asymptomatic women of gestational age (Brown et al.

✉ Navid Nezafat
n.nezafat@srbiau.ac.ir

✉ Amir Savardashtaki
Dashtaki63@gmail.com

¹ Department of Medical Biotechnology, School of Advanced Medical Sciences and Technologies, Shiraz University of Medical Sciences, Shiraz, Iran

² Pharmaceutical Sciences Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

³ Department of Pharmaceutical Biotechnology, School of Pharmacy, Shiraz University of Medical Sciences, Shiraz, Iran

⁴ Department of Parasitology & Mycology, School of Medicines, Shiraz University of Medical Sciences, Shiraz, Iran

⁵ Department of Pathology, Shiraz University of Medical Sciences, Shiraz, Iran

⁶ Infertility Research Center, Shiraz University of Medical Sciences, Shiraz, Iran

2019; Gharaghani et al. 2018). Azole antifungal agents such as miconazole, ketoconazole, and fluconazole are effective treatments for *Candida albicans* infections. Nevertheless, increasing drug resistance has made this treatment problematic. Therefore, the early detection of fungal infections is impressive for treatment, prevention of drugs overuse, and reduction of costs (Thomson et al. 2015).

Various identification methods such as culture on Corn Mill Agar-Tween 80, sugar uptake test by UX 20C API kit, and Germ tube test have been used to identify *Candida* species (Souza et al. 2015). Though blood culture is the standard method for diagnosing candidiasis, for the low sensitivity of the method (about 50%), positive results appear only in the advanced stages of infection, which significantly impairs the treatment process. Although molecular methods such as Polymerase Chain Reaction (PCR), cell wall, and cytoplasm antigens identification are used to early detect the disease, none of these methods has been widely approved (He et al. 2015). Immunological methods like Immunofluorescence Antibody Test (IFAT), Enzyme-linked Immunoassay (ELISA), and immunoblotting are more sensitive than molecular methods and help assess the host's immune response (Buonfrate et al. 2015; Levenhagen and Costa-Cruz 2014). Considering the production of recombinant antigens and synthetic peptides that have immunogenic epitope domains, desirable improvement has been achieved in serological methods (Sefidi et al. 2013). Extensive efforts have been made to establish a valid serological test to diagnose invasive candidiasis. Detecting specific antibodies against yeast antigens in patient samples is one of the documented methods (Calandra et al. 2016; Ma et al. 2013). Based on the results of several studies, various proteins from *Candida albicans* were regarded as antigens. Proteins that stimulate the immune system and produce antibodies are found in two categories. The first group contains proteins that appear specifically in the invasive phase of the disease, such as HYR1, HWP1, ECE1, and ALS. The expression of the second group, including EAP1, SAP1, BGL2, and MET6, increases in the pathogenic stage (Araújo et al. 2017; Laín et al. 2008; Mayer et al. 2013; Wang et al. 2015).

The last decade has witnessed a considerable increase in the processing speed of computers and the storage capacity of information. On the one hand, sequencing the genomes of many prokaryotes, eukaryotes, and many proteins gives new insights into biological data analysis (Bansal 2005). In addition, the design, production, and purification of each of the essential antigens for the diagnosis of infectious diseases are time-consuming, requiring high costs and controlling different processes. Thus, combining these various antigenic proteins as a chimeric protein in a vector can be a valuable way to reduce the time and cost of processing. This study aimed to bioinformatically design and clone the recombinant

Candida albicans multi-epitope antigen to diagnose invasive candidiasis using the ELISA system.

Methods

Study Design

The current study was done in three following steps. In the first step, according to the results of *Candida albicans* proteomic microarray and literature review on *Candida albicans* antigens, the antigenic proteins of *Candida albicans* were chosen. Invasive phase-specific proteins, including HWP1, ALS, ECE1, and HYR1, and proteins associated with incremental expressions of the pathogenic stage, such as BGL2, EAP1, MET6, and SAP1, were the select antigens. In the next step, the above-mentioned antigens from different *Candida albicans* strains were aligned via CLC Genomics Workbench 20 to select the consensus regions. Next, using immunoinformatics servers, the antigenic regions and B cell epitopes of each protein were defined, and then the final epitopes were joined by an appropriate linker. Afterward, the structural analysis of the final construct, such as 3D modeling, physicochemical properties, and validation, was performed. In the ultimate step, the protein sequence was reverse-translated to gene sequence and optimized for expression in the *Escherichia coli* (*E. coli*) host.

Sequence Retrieval

The amino acid sequences of BGL2, EAP1, MET6, SAP1, HWP1, ALS, ECE1, and HYR1 were retrieved from the Protein knowledgebase (UniProtKB) database (www.uniprot.org) in the FASTA format. Also, CLC Genomics Workbench 20 was applied for antigenic protein aligning. The reviewed sequences of ALS protein were aligned to achieve a protected sequence of this protein.

Immunoinformatics Analyses

Antigenicity Prediction

The EMBOSS server 6.6.0.0 (<https://www.bioinformatics.nl/cgi-bin/emboss/antigenic>) was used to define antigenic regions of the proteins. The performance of this server is based on the Kolaskar and Tongaonkar technique, which forecasts protein antigenic regions using physicochemical properties of amino acid residues. In the present study, initial parameters such as antigenic region with a minimum length of 6 and output report format EMBOSS motif were utilized to establish the antigenic region of the target proteins (Rice et al. 2000).

Prediction of Linear B Cell Epitopes

B cell epitopes considerably affect antibody production. For this purpose, in this study, B cell epitopes were determined using three servers to increase the prediction accuracy. In the beginning, for predicting linear B cell epitopes, ABCpred (https://webs.iitd.edu.in/raghava/abcpred/ABC_submission.html) was utilized to predict linear B cell epitopes in the antigenic sequence using an artificial neural network algorithm. In addition, ABCpred is the first server developed according to the recurrent neural network by fixed-length samples with about 65.93% accuracy (Chatterjee et al. 2021; Saha and Raghava 2006).

The BCPRED 2.0 server (https://webs.iitd.edu.in/raghava/bcpred/bcpred_submission.html) was also applied to predict linear B cell epitopes. This server predicts B cell considering the physicochemical properties, including accessibility, hydrophilicity, polarity, flexibility/mobility, exposed surface, turns, or a combination of these features. Based on different physicochemical properties, the server can predict epitopes with various accuracies from 52.92 to 57.53%. In the case of using flexibility, hydrophilicity, polarity, and surface features simultaneously at a threshold of 2.38, this server can predict epitopes with 58.7% accuracy (Saha and Raghava 2004).

BepiPred-2.0 (<http://www.cbs.dtu.dk/services/BepiPred/>) is another server that predicts B cell epitopes based on epitope information from solved 3D structures and a wide range of linear epitopes downloaded from the IEDB database. About 50 sequences can be submitted simultaneously, with a length of 10–6000 amino acids. The prediction for each sequence usually takes a few minutes (Jespersen et al. 2017).

Epitopes Antigenicity Prediction

The Vaxijen 2.0 server (freely available online at <http://www.jenner.ac.uk/VaxiJen>) was used to analyze the antigenicity of each epitope in a multi-epitope protein structure. The VaxiJen v2.0's threshold value was set at 0.5. The prediction method on this server is alignment-free and based on the protein's various physicochemical properties.

Epitopes Allergenicity Prediction

The AllerTOP V2.0 server (<http://www.ddg-pharmfac.net/AllerTOP>) was used to predict the allergenicity of the designed construct's epitopes. To classify allergens, the server employs methods such as amino acid E-descriptors, auto- and cross-covariance transformation, and k nearest neighbors (kNN) machine learning (Dey et al. 2022).

Multi-epitope Candidate Construction

The B cells epitope candidates, which were selected by relevant servers and got the highest scores, were ligated using a suitable amino acid linker. Eventually, an amino acid structure with multiple epitope characteristics was constructed.

Physicochemical Properties Assessment

The ProtParam tool at the EXPASY server (<http://web.expasy.org/protparam/>), which is the proteomics server of the Swiss Institute of Bioinformatics, was utilized for calculating different physicochemical features of proteins. The parameters consisted of theoretical pI (isoelectric point), molecular weight (kilodalton), instability index, estimated aliphatic index, and grand average of hydrophobicity (GRAVY) (Gasteiger et al. 2005).

Tertiary Structure Prediction of Designed Construct

Galaxy TBM (<http://galaxy.seoklab.org/cgi-bin/submit.cgi?type=TBM>) was applied to predict the tertiary protein structures. Galaxy TBM performs homology modeling using comparative analysis of the amino acid sequence with those of proteins whose structures are experimentally determined to predict the structure of the regions of proteins that do not have a high degree of homology with other proteins (Ko et al. 2012).

Tertiary Structure Validation

The ProSA-web server (<https://prosa.services.came.sbg.ac.at/prosa.php>) is one of the servers used for evaluating the 3D structure of the protein. The software mainly revolves around the alpha carbon of amino acid residues to determine the atomic interactions of the protein query. The energy level of the query structure, depicted via a color code system, determines the z-score, which illustrates the quality of the protein structure. Natively folded proteins lie within a defined z-score value, which is used as the reference range with which query proteins are evaluated. The ProSA-web server is accessible elsewhere (Wiederstein and Sippl 2007).

The conformation of proteins is defined by the ϕ and ψ values, which are the allowable freedom of rotation in C α -N and C α -C of peptide bonds, respectively. The allowable torsion angles between atoms are those that do not cause any steric hindrance; thus, the spherical shape and van der Waals radii were meticulously analyzed. The models were categorized into three groups using the above

parameters: Favored, allowed, and outlier regions. The plot was generated using PROCHECK online software (<https://saves.mbi.ucla.edu/>) (Lovell et al. 2003).

The ERRAT online software (<http://services.mbi.ucla.edu/ERRAT/>) was employed to verify the 3D structure of proteins. The program uses non-covalently bonded atom–atom interactions and determines whether regions of the target proteins are correctly folded or not. Six types of atomic interactions, including C–C, C–N, C–O, N–N, and N–O, are evaluated where C, N, and O denote carbon, nitrogen, and oxygen, respectively. The software posits that the distance between the aforementioned non-bonded interaction should not surpass 3.5 Å, and the covalently bonded atoms within the same polypeptide strand should be excluded from the analysis (Colovos and Yeates 1993).

Molecular Dynamics (MD) Simulations

Understanding macromolecular functioning requires dynamic systems simulations with biologically relevant sizes and time scales. As an alternative to atomistic simulations, coarse-grained representations combined with Normal Mode Analysis (NMA) have been established. Internal Coordinates (IC) can significantly improve the versatility and efficiency of current approaches that are usually based on Cartesian coordinates. The iMODS server (<https://imods.iqfr.csic.es/>) was employed to explain protein collective motion (López-Blanco et al. 2011). To investigate the motions of proteins using NMA in internal coordinates, one can simply submit the PDB-ID or the atomic coordinates in PDB format.

In Silico Cloning

Codon optimization and back-translation of the final construct were done by the OPTIMIZER server (<http://genom.es.urv.es/OPTIMIZER/>) for the high-level expression of the multi-epitopic protein structure in *E. coli*. Afterward, the different factors of DNA sequence such as Codon with a Frequency Distribution (CFD), overall GC content, and no negative CIS elements and repeat sequences were evaluated by GenScript Rare Codon Analysis Tool (<https://www.genscript.com/tools/rare-codon-analysis>). In the next step, the pET-23a+ expression vector was used to simulate in silico cloning of the DNA sequence of the construct.

Results

Sequence Retrieval

The UniprotKB database was used to obtain the amino acid sequences of eight pathogenic *Candida albicans* proteins. All sequences of MET6, two sequences of BGL2, three

sequences of EAP, 17 sequences of SAP1, 24 sequences of HWP1, 44 sequences of ALS, four sequences of ECE1, and 21 sequences of HYR1 were retrieved from UniProtKB (Table 1).

Immunoinformatics Analyses

Antigenicity Prediction

The EMBOSS server was used to detect the antigenic sequence with the highest score for antigenicity. Approximately 13 antigenic regions of BGL2 protein, 27 antigenic regions of EAP1 protein, 32 antigenic regions of MET6 protein, 18 antigenic regions of SAP1 protein, 25 antigenic regions of HWP1 protein, 23 antigenic regions of ALS protein, 13 antigenic regions of ECE1 protein, and 32 antigenic regions of HYR1 protein were identified. All the results obtained from this server were found in Supplementary Tables (S1a–h).

Determining Linear B Cell Epitopes

The ABCpred, BCPRED, and BepiPred servers were used to detect linear B-cell epitopes from all eight proteins based on their amino acid sequences. The number of linear epitopes and their scores are listed in Table 2.

The shared sequences predicted by the two mentioned servers were selected by Microsoft Excel software, and the best immunogenic epitopes were chosen. The finally selected regions of BGL2, EAP1, MET6, SAP1, HWP1, ALS, ECE1, and HYR1 proteins are presented in Table 3.

Antigenicity and Allergenicity of Designed Construct's Epitopes

The antigenicity of each epitope was estimated separately using the VaxiJen 2.0 server and a fungal model. Epitopes were identified as antigens or potential antigens based on the findings. The scores obtained from the VaxiJen 2.0 server are as follows: BGL2 is 1.35, EAP1 is 1.26, MET6 is 0.5, SAP1 is 0.76, HWP1 is 0.96, ALS is 0.93, ECE1 is 0.56, and HYR1 is 1.67. Allergenicity prediction by the AllerTOP server showed that all epitopes in the designed structure were non-allergenic.

Assembling Epitopes

The select B cell epitopes of BGL2, EAP1, MET6, SAP1, HWP1, ALS, ECE1, and HYR1 proteins, displayed in Table 3, were connected using an EAAAK linker. Linkers are short peptides to connect different protein domains. Building a multi-epitope diagnostic kit without a proper linker can exacerbate protein folding issues. The EAAAK

Table 1 The list of epitope sequence of BGL2, EAP1, MET6, SAP1, HWP1, ALS, ECE1 and HYR1 proteins obtained from UniprotKB

Name of protein	Accession number	Length of amino acid (aa)
BGL2	Q5AMT2	308
BGL2	P43070	308
EAP1	G1UBC2	653
EAP1	Q5AP80	785
EAP1	Q5AQ12	176
MET6	P82610	767
SAP1	P0CY27	391
SAP1	P0DJ06	398
SAP1	Q5AC08	418
SAP1	P0CY29	398
SAP1	Q5A8N2	417
SAP1	Q59VH7	588
SAP1	Q5A651	453
SAP1	Q5AEM6	405
SAP1	P0CS83	398
SAP1	C4YMJ3	398
SAP1	C4YSF6	391
SAP1	C4YNQ5	398
SAP1	O42778	405
HWP1	C4YHA6	633
HWP1	Q59PF9	908
HWP1	Q00312	527
HWP1	Q5ABZ2	534
HWP1	Q5AF41	213
HWP1	Q5AF39	113
HWP1	Q59TP1	721
HWP1	Q07730	271
HWP1	Q5A312	527
HWP1	Q5AAF4	1485
HWP1	Q5A220	600
HWP1	Q5AL52	1732
HWP1	Q5AL03	919
HWP1	Q5A0E5	310
HWP1	Q59U10	740
HWP1	P0CY34	512
HWP1	Q5AMQ6	1111
HWP1	Q5A8T4	1260
HWP1	Q59L12	1155
HWP1	Q5A8T7	1347
HWP1	A0A1D8PQ86	1890
HWP1	A0A1D8PQB9	2100
HWP1	P0CU38	2530
ALS	Q59L12	1155
ALS	Q5A312	1568
ALS	Q5A8T7	1347
ALS	A0A1D8PQ86	1890
ALS	A0A1D8PQB9	2100

Table 1 (continued)

Name of protein	Accession number	Length of amino acid (aa)
ALS	Q5A2Z7	1366
ALS	Q5A8T4	1260
ALS	Q9URQ0	874
ALS	P0CU38	2530
ALS	O13368	1419
ALS	O74660	875
ALS	O74623	1119
ALS	P46590	1260
ALS	Q59U10	740
ALS	Q9P940	248
ALS	Q5APJ9	219
ALS	P46593	634
ALS	Q5A0E5	310
ALS	O74189	877
ALS	Q5AMQ6	1111
ALS	Q5A220	600
ECE1	Q07730	271
ECE1	P46593	634
ECE1	Q5A0E5	310
ECE1	Q5AMQ6	1111
ECE1	Q5A220	600
ECE1	Q8X3T7	29
HYR1	Q5AL03	919
HYR1	Q59U10	740
HYR1	P87020	299
HYR1	Q59XX2	378
HYR1	P46593	634
HYR1	Q5A0E5	310
HYR1	Q5AMQ6	1111
HYR1	Q5A220	600
HYR1	Q5APU2	597
HYR1	Q59XA7	1249
HYR1	Q5A849	1225
HYR1	Q59XL0	1085
HYR1	Q5AAL9	1526
HYR1	Q5A7R7	511
HYR1	P46591	937
HYR1	Q59XB0	714
HYR1	Q5A1E0	1308
HYR1	Q5A5M7	1562
HYR1	Q5A6U1	941
HYR1	Q5A029	941
HYR1	Q59L86	872

rigid linker was used to assemble the designed diagnostic kit. The EAAAK linker not only maintains proper spacing between epitopes and fragments of the designed vaccine but also eliminates unwanted interactions. Also, a 6xHis

Table 2 Linear B cells epitopes analysis of BGL2, EAP1, MET6, SAP1, HWP1, ALS, ECE1 and HYR1 proteins by ABCpred, BCPRED and BepiPred servers

Protein	Server	Number of linear B cell epitopes	Maximum score
BGL2	ABCpred	29	0.91
BGL2	BCPRED	6	0.999
BGL2	BepiPred	11	0.66
EAP1	ABCpred	84	0.95
EAP1	BCPRED	26	1
EAP1	BepiPred	1	0.66
MET6	ABCpred	78	0.97
MET6	BCPRED	10	0.999
MET6	BepiPred	28	0.67
SAP1	ABCpred	33	0.94
SAP1	BCPRED	7	0.997
SAP1	BepiPred	11	0.7
HWP1	ABCpred	65	0.97
HWP1	BCPRED	26	1
HWP1	BepiPred	1	0.71
ALS	ABCpred	62	0.95
ALS	BCPRED	19	1
ALS	BepiPred	15	0.65
ECE1	ABCpred	24	0.90
ECE1	BCPRED	3	1
ECE1	BepiPred	7	0.65
HYR1	ABCpred	92	0.92
HYR1	BCPRED	27	1
HYR1	BepiPred	8	0.7

tag was added to the C-terminus to facilitate the purification process of multi-epitope protein structures (Fig. 1).

Assessment of Physicochemical Properties

Different physicochemical properties of the designed multi-epitope construct were determined via the ProtParam server. The construct's number of amino acids and molecular weight

(MW) were 255 and 27.3 kDa, respectively. The theoretical pI of the construct to indicate the basic condition of the protein was 6.80. The estimated half-life was 2.8 h in mammalian reticulocytes, 3 min in yeast, and 2 min in *E. coli*. The aliphatic index was 71.37, and the grand average of hydropathicity (GRAVY) was -0.289 , which showed that the protein was hydrophobic and could better interact with neighboring water molecules. The instability index was computed as 32.48, which classified the protein as stable.

Tertiary Structure Modeling

The 3D model of the structure was predicted by the GalaxyTBM server, resulting in five initial models (Fig. 2a–e).

Tertiary Structure Validation

The ERRAT (Fig. 3a–e), ProSA-web (Fig. 4a–e), and Ramachandran plot servers (Fig. 5a–e) were used to evaluate the quality of the predicted models to select the best model. The ProSA-web server shows the overall quality score of protein (z -score) obtained by comparing the 3D model of the target protein with similar proteins of the same size deposited in PDB. The validation results of the five modeled structures by the ERRAT, ProSA-Web and Ramachandran servers are presented in Table 4. After comparing all the obtained results from the servers mentioned above, model 4 was selected as the best structure. In this model, the ERRAT quality and ProSA z -score were 78.59 and -3.33 , respectively (Table 4). Moreover, the Ramachandran plot showed that 86.6%, 11.5%, and 1.5% of the amino acid residues were located in the favored region, allowed region, and generously allowed regions, respectively.

Molecular Dynamics (MD) Simulations

Understanding the macromolecular function of the protein complex was performed using the internal coordinates of the iMODS server (Fig. 6a). The complex's eigenvalue was

Table 3 The final selected B cell epitopes regions of BGL2, EAP1, MET6, SAP1, HWP1, ALS, ECE1 and HYR1 proteins

Protein	Start–end position	Sequence	Length of sequence
BGL2	276–306	WKPDTSGTSSVEKHGWVWQSDKTLKYSIDCK	31
EAP1	91–114	STSYTTDTAYTTTIVITVTKCDGGS	24
MET6	77–110	NAIPERYTKFDLAPIDVLFAMGRGLQKKATETQA	34
SAP1	22–53	KRSPGFVTLDFDVIKTPVNATGQEGKVKRQAI	32
HWP1	376–408	TTVITVTSCESSCTESEVTTGVVVVVTSEETVY	33
ALS	313–354	TGYRNSDAGSNGIVIVATTRTVTDSTTAVTTLFPDPNRDKTK	42
ECE1	79–103	QIIMSIVKAFKGNKREDIDSVVAGII	25
HYR1	831–863	ANPVTSTESDTTISVTVSITSYMTGFDGKPKP	33

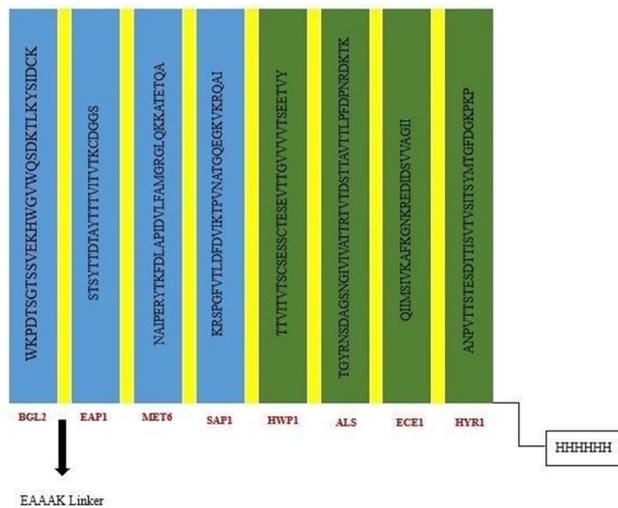


Fig. 1 Schematic diagram of assembled multi-epitope structure. B cell epitopes were selected from antigenic proteins (BGL2, EAP1, MET6, SAP1, HWP1, ALS, ECE1 and HYR1). Using EMBOSS, BepiPred, ABCpred, and BCPRED servers. The selected sequences were connected to each other the utilization EAAAK linkers. as well as, a 6xHis tag was added to the C-terminus for ease of purification

2.744204e40 (Fig. 6b). Arrows indicated the direction of each residue in the 3D model, and the length of the line represented the extent of mobility. The B-factor values derived from normal mode analysis were comparable to RMS (Fig. 6c). Hinges in the chain indicated the complex's probable deformability as measured by the contortion of each individual residue (Fig. 6d). Each normal mode's variance was inversely related to its eigenvalue. The covariance matrix explained the link between pairs of residues by representing correlated, uncorrelated, or anti-correlated motions with red, white, and blue colors, respectively (Fig. 6e). As a result, an elastic network model (Fig. 6f) was generated, which identified the pairs of atoms connected by springs. Each dot in the diagram represented one spring between the corresponding pair of atoms and was colored according to its stiffness.

In Silico Cloning of the Final Construct

The OPTIMIZER online server was used for back-translation and codon optimization of the multi-epitope construct for this study. The gene sequence's CAI, CFD, and GC content were 0.88, 0%, and 53.3%, respectively (Fig. 7a–c). A CAI of > 0.8 is considered good for expression in a selected host. The optimal percentage range of GC content is between 30 and 70%. Codons with a frequency distribution (CFD) < 30% are likely to impede transcriptional and translational efficiency. All the above-mentioned factors of our sequence were in the ideal range. This indicated that the final multi-epitope construct with optimized sequence

would be efficiently expressed in *E. coli*. Next, an optimized nucleotide sequence was cloned into the pET-23a+ expression vector. For this purpose, *NdeI* and *XhoI* restriction sites were added to the N and C-terminals of the sequence, respectively (Fig. 8).

Discussion

The occurrence of invasive candidiasis was increased by 50% globally over the past decade, varying from 2.4 per 100,000 to 15 per 100,000, according to the clinical and national conditions. Despite considerable advances in antifungal remedies, candidiasis spreads with high mortality (40–60%). The timely use of appropriate antifungal agents is necessary to increase survival. However, yeast blood cultures are much less available and require long-term (24 h or longer) incubation (Poissy et al. 2020). ELISA and concomitant immunological methods are used in most detection approaches due to desirable features such as high sensitivity, convenience, and relatively low cost (Khatami et al. 2020b). Recombinant antigenic proteins are the best options for detecting antibodies against invasive *Candida* antigenic proteins in the blood. Also, as an efficient serological diagnostic method, ELISA is one of the most effective procedures for detecting antibodies in invasive candidiasis. Using the complete sequence of an antigen to detect antibodies can cause nonspecific binding to the target antibody due to the similarity and overlap with similar sequences of other living antigenic proteins. In general, this method reduces the degree of specificity (Dey et al. 2021; Laín et al. 2007b). Each antigen's design, production, and purification in diagnosing infectious diseases are time-consuming, requiring high costs and controlling different processes. Thus, a combination of different antigenic proteins as a chimeric protein can be a valuable way to reduce the time and cost of the process. In recent years, bioinformatics knowledge has increased in various scientific fields, e.g., the production of recombinant products and proteins (Khatami et al. 2020a; Mahapatra et al. 2020; Sabetian et al. 2019; Shahbazi et al. 2016; Tehrani et al. 2020; Zarei et al. 2019). The present study used various bioinformatics strategies to design the *Candida albicans* diagnostic kit based on linear B cell epitopes of BGL2, EAP1, MET6, SAP1, HWP1 ALS, ECE1, and HYR1 antigens. Zheng et al. conducted serological assays for rapid diagnosis of invasive candidiasis. In this regard, recombinant BGL2 protein of *Candida albicans* cell wall was produced, and an indirect ELISA method was developed to detect IgG antibodies. The sensitivity, specificity, positivity, and negative predictive values were 80.8%, 90.0%, 89.4%, and 81.8%, respectively for anti-BGL2 detection. The results showed that recombinant BGL2 antigen from *Candida albicans* reacted to the sera of proven IC patients infected with *Candida albicans*

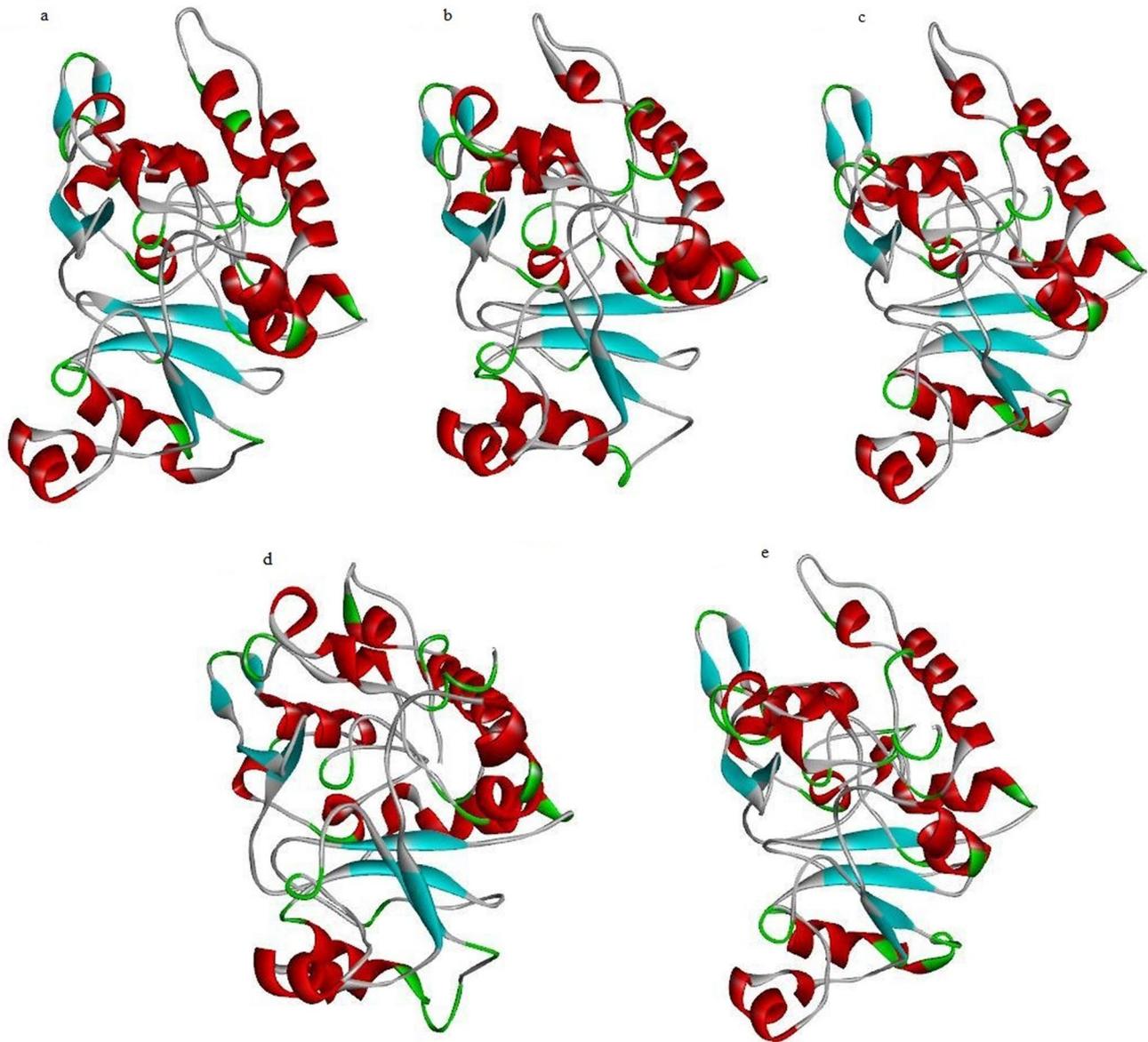


Fig. 2 In order to predict the third structure of the designed multi-epitope structure, GalaxyTBM server was used. As a result, five different models of designed multi-epitope protein by GalaxyTBM server, **a** model 1, **b** model 2, **c** model 3, **d** model 4, **e** model 5

(He et al. 2015). The EAP1 protein, expressed by the EAP1 gene, binds to glycosylphosphatidylinositol in the *Candida albicans* cell wall. The EAP1 protein is involved in surface adhesion. Different studies have shown that mutations in this protein reduce *Candida albicans* adhesion to epithelial cells, indicating the importance of this protein in the binding of yeast cells. The EAP1 protein is also effective in forming and developing *Candida albicans* biofilms (Li et al. 2007). Recently, it was discovered that EAP1 plays a role in macrophage regulation and lymphocyte activation. Furthermore, EAP1 suppressed CD69 expression, implying

that EAP1 is involved in CD4T-cell activation (Ramalho-Oliveira et al. 2019). Secreted Aspartic Proteinases (SAPs) are a group of hydrolytic enzymes secreted by *Candida albicans* to penetrate host cell membranes. There is an unequal tendency for SAPs proteins to penetrate various cell types. Therefore, in different stages of yeast infection, they could be expressed and secreted into both intracellular and extracellular environments. In this regard, the SAP1 protein is mainly observed in vaginal infections caused by *Candida albicans* (Oliver et al. 2019; Singh et al. 2019). According to studies, the increase in Th1/Th2/Th17 cytokine levels

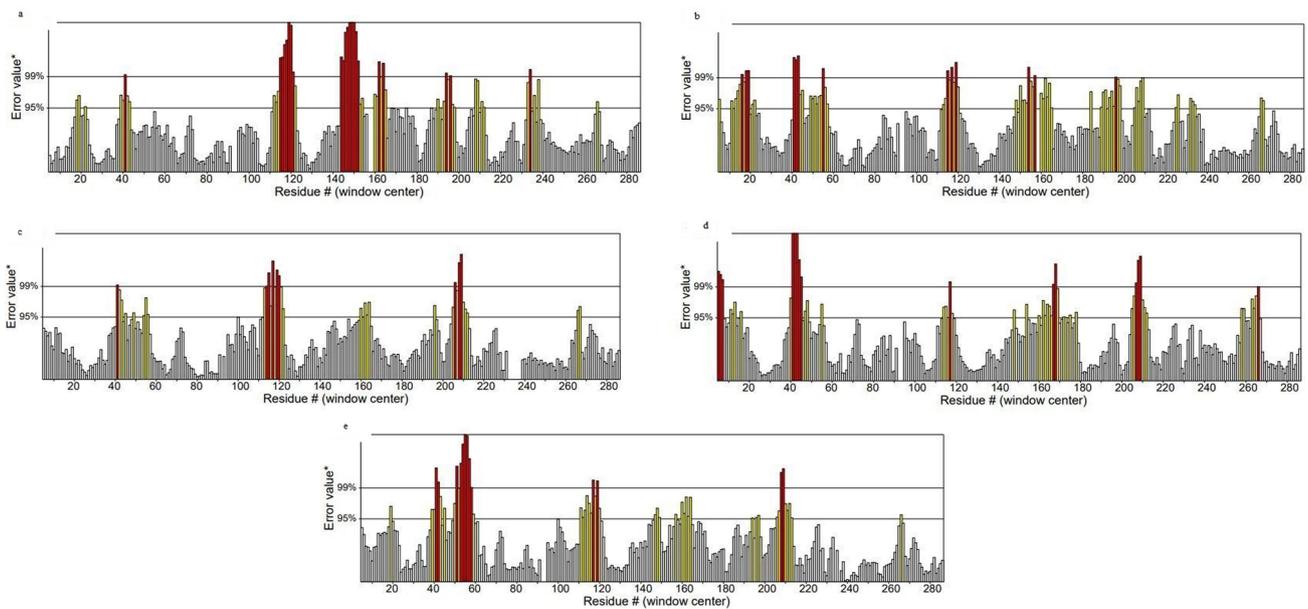


Fig. 3 The 3D structure validation results of multi-epitopic protein construct with ERRAT plot. The overall quality factor obtained from the ERRAT server for the proposed models are as follows: **a** Model

1 (78.4), **b** Model 2 (75.3), **c** Model 3 (78.6), **d** Model 4 (77) and **e** Model 5 (76.5)

observed in SAP2-vaccinated mice indicates that SAP2 has immunomodulatory properties. SAP2 immunization also significantly increases total CD45+ leukocytes in the spleen, preventing a significant decrease in their number following fungal infection. Furthermore, SAP2 immunization increases plasma cell counts and the proportion of fungal-binding B cells in the spleens of immunized mice (Shukla and Rohatgi 2020). Leighann et al. showed that using a defined panel of genes during the initial growth of biofilm may be very significant for diagnosing *Candida albicans*. Panels of genes such as SAP, HWP1, and EAP1 are examined sequentially to identify different biomarkers for diagnosing invasive candidiasis in immunocompromised patients (Sherry et al. 2014). Methionine synthase (MET6) is a protein with an essential role in methionine biosynthesis during *Candida albicans* morphogenesis (Sáez-Rosón et al. 2014). Patients with disseminated candidiasis have a better prognosis when MET6-specific IgG serum antibodies are produced (Adams et al. 2021). Hyphal Wall Protein 1 (HWP1) facilitates *Candida albicans* attachment to host epithelial cells. In addition, HWP1 encodes a principal protein of *Candida albicans* that is involved in cell wall proliferation and intracellular signaling (Orsi et al. 2014). Agglutinin-Like Sequence (ALS) proteins are important proteins in the *Candida albicans* cell wall, which encode large cell surface glycoproteins and play a key role in adhesion to different surfaces (Lombardi et al. 2019). Generally, a valid diagnostic method is used to detect antibodies against ALS, HWP1, and MET6 *Candida albicans* proteins. The results obtained by ELISA showed

that the sensitivity to detect the desired antibodies is more than 66%. Also, in immunocompromised patients, anti-MET6 antibodies were named the best biomarkers (Díez et al. 2021). ECE1p is a large protein with 271 aa that is processed in the Endoplasmic Reticulum (ER) and the Golgi complex. Candidalysin toxin is also derived from the Ece1p protein (Richardson et al. 2018). Antifungal drugs have been shown to eliminate various components of fungi required for their virulence. They also influence the expression of genes involved in hyphae growth, such as Hwpl and Ecel, in the epithelium. As a result, Hwpl reduced IL-8 production, inhibited aspartyl proteinase expression, and reduced the inflammatory response (Rodríguez-Cerdeira et al. 2020). A Hyphally Regulated Gene (HYR1) is expressed specifically during hyphal development. Hyr1p contributes to *Candida albicans* virulence by resisting phagocyte killing (a major host defense mechanism against candidiasis) via an unknown mechanism. (Uppuluri et al. 2018). Numerous studies have investigated the recombinant *Candida albicans* antigens for their potential use in diagnosing invasive candidiasis. In this context, SAP, HWP1, HYR1, ECE1, and ALS proteins as *Candida albicans* recombinant antigens have been tested for serodiagnosis of invasive candidiasis (Lafn et al. 2007a, 2008).

In the current study, the proteins sequences were obtained from the UniProt databases at www.uniprot.org. Prediction of B cell epitopes plays a vital role in developing diagnostic kits based on multi-epitope proteins. Due to the subject's significance, three different servers were used to identify the

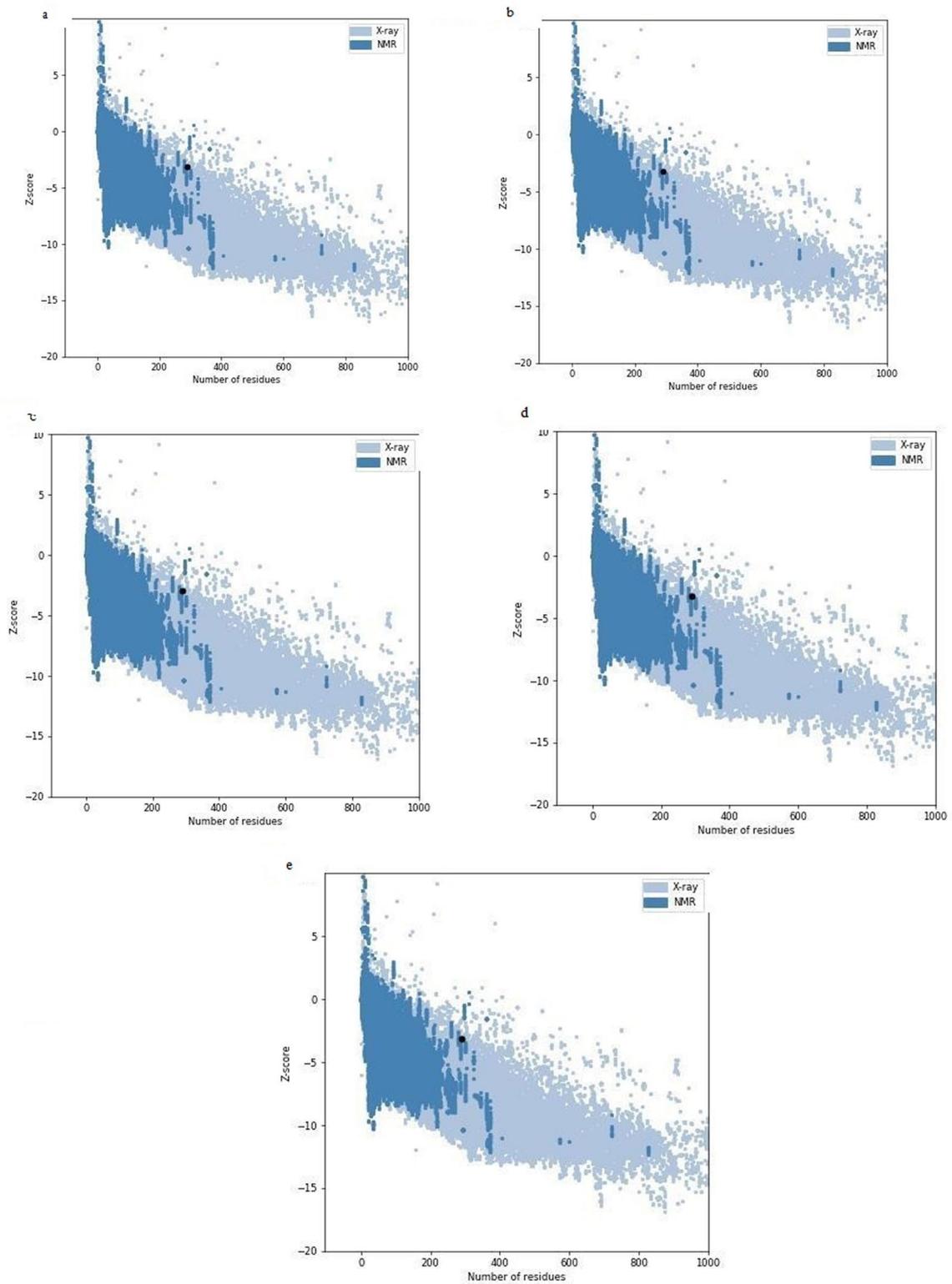


Fig. 4 ProSA-Web z-score, indicate the validation of protein 3D models. The ProSA Z-score, based on X-ray crystallography and NMR spectroscopy, of Models **a** 1, **b** 2, **c** 3, **d** 4, **e** 5, are -3.33 , -3.23 , -2.9 , -2.96 , and -3.4 , respectively

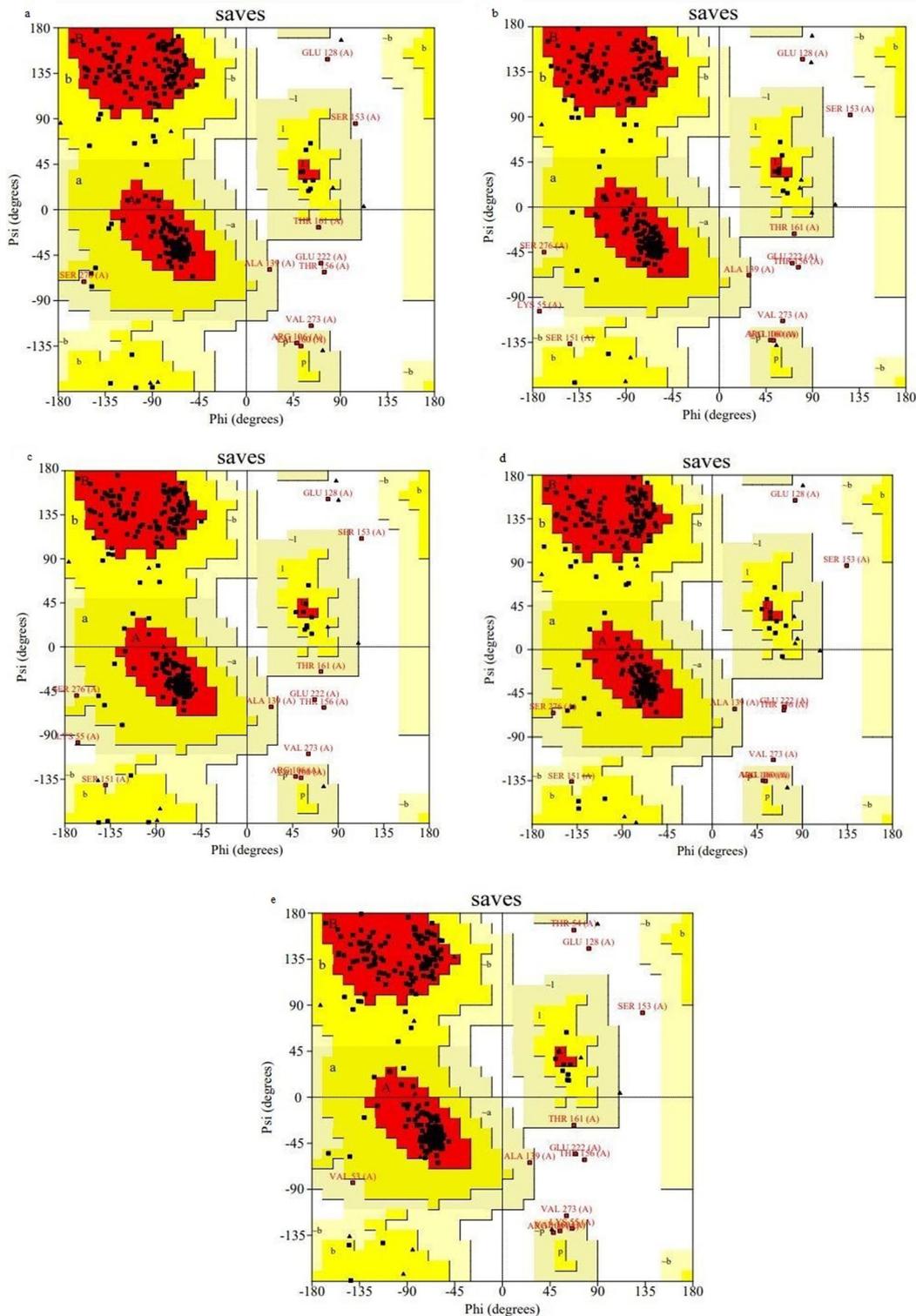


Fig. 5 The PROCHECK's Ramachandran plot validate the 3D models of selected multi-epitopic construct. Favoured regions in models was shown; **a** model 1 (87.0%), **b** model 2 (87.4%), **c** model 3 (86.6%), **d** models 4 (85.9%), **e** model 5 (86.3%)

linear 16-mer B cell epitopes regions. The final epitopes were selected and then connected by the EAAAK linker. Selecting a suitable linker is crucial for the appropriate binding of protein

fragments in a multi-epitope structure. The linkers are responsible for connecting protein segments and have an essential role in maintaining the biological activity and increasing the

Table 4 The ERRAT, ProSA-Web AND Ramachandran plot servers analysis

Models	Ramachandran plot results	ERRAT overall quality factor	ProSA-Web (z-scores)
Model 1	Favoured regions: 87.0% Allowed regions: 11.1% Generously allowed regions: 3 1.1% Disallowed regions: 0.8%	78.4	−3.33
Model 2	Favoured regions: 87.4% Allowed regions: 10.03% Generously allowed regions: 1.5% Disallowed regions: 0.8%	75.3	−3.23
Model 3	Favoured regions: 86.6% Allowed regions: 11.5% Generously allowed regions: 1.5% Disallowed regions: 0.4%	78.6	−2.9
Model 4	Favoured regions: 85.9% Allowed regions: 11.8% Generously allowed regions: 1.5% Disallowed regions: 0.8%	77	−2.96
Model 5	Favoured regions: 86.3% Allowed regions: 11.5% Generously allowed regions: 1.5% Disallowed regions: 0.8%	76.5	−3.4

expression of the structure (Chen et al. 2013; Li et al. 2016; Mahapatra et al. 2021). The EAAAK (Glu–Ala–Ala–Ala–Lys) sequence is a tough and helix-forming linker. Lee et al. conducted a study to enhance the antibacterial activity of an attacin-coleopteracin hybrid protein fused with a linker. They used helical (Glu–Ala–Ala–Ala–Lys) and non-helical linkers (Gly–Gly–Gly–Gly–Ser) to compare them in binding the two antimicrobial proteins. The results showed that the fused proteins with the EAAAK linker had more antibacterial properties (Lee et al. 2013). Our study selected and ligated antigenic regions of different proteins using linkers. The three-dimensional structures of the proteins were modeled and evaluated by several servers. In this way, the Galaxy-TBM server developed five models for the desired structure. Then, the obtained models were evaluated by the ERRAT, ProSA-Web, and Ramachandran plot servers, and finally, model 4 was selected.

Conclusion

In this present study, to provide a diagnostic kit for *Candida albicans*, we used immunodominant linear B cell epitopes from eight pathogenic *Candida albicans* antigens (BGL2, EAP1, MET6, SAP1, HWP1, ALS, ECE1, and HYR1) to model and design a multi-epitope protein using suitable linkers. The structural and physicochemical evaluations of the designed protein construct showed that our recombinant protein was thermostable, hydrophilic, and basic. Finally, it was found that the designed chimeric protein could be significantly expressed in *E. coli*. The multi-epitope-designed protein can be used to construct a high-performance serological test for the diagnosis of candidiasis infections.

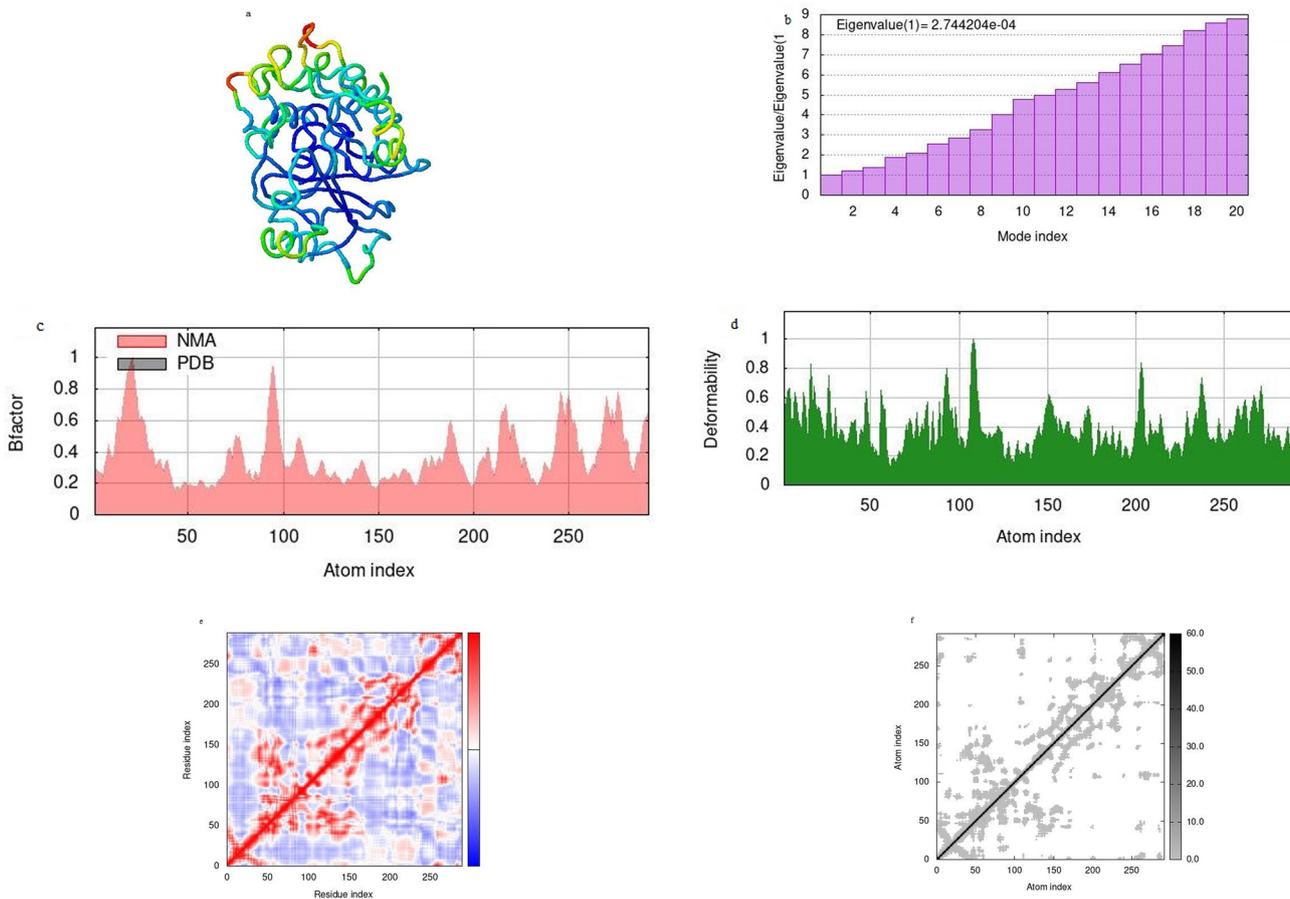


Fig. 6 Molecular dynamics simulation of multi-epitope diagnostic kit; stability of the protein complex was **a** investigated through mobility, **b** eigenvalue, **c** B-factor, **d** deformability, **e** covariance and **f** elastic network analysis

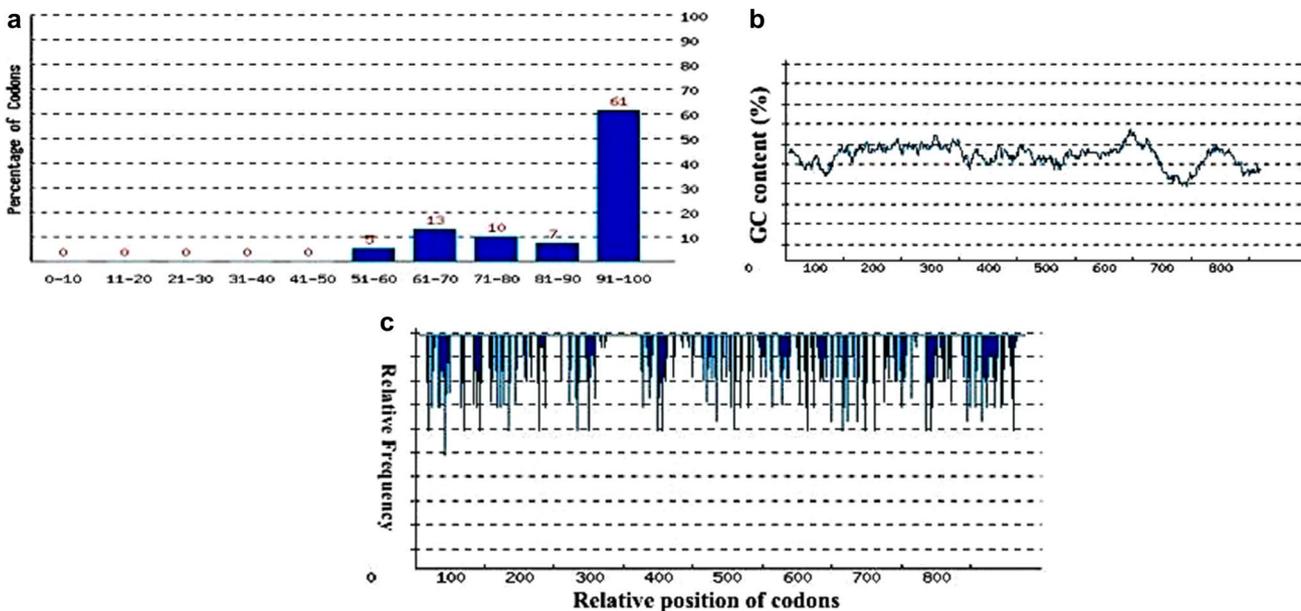


Fig. 7 Designed structural analysis to express high levels of multi-epitope protein in *E. coli* host. **a** Codon frequency distribution (CFD) value of gene sequence is 91–100, 81–90, 71–80, 61–70, 51–60 in the

gene is 61%, 7%, 10%, 13%, and 5%, respectively. **b** The average GC content of our sequence is 53.3%. **c** CAI of gene sequence is 0.88; a CAI of 40.8 is considered as good for expression in selected host

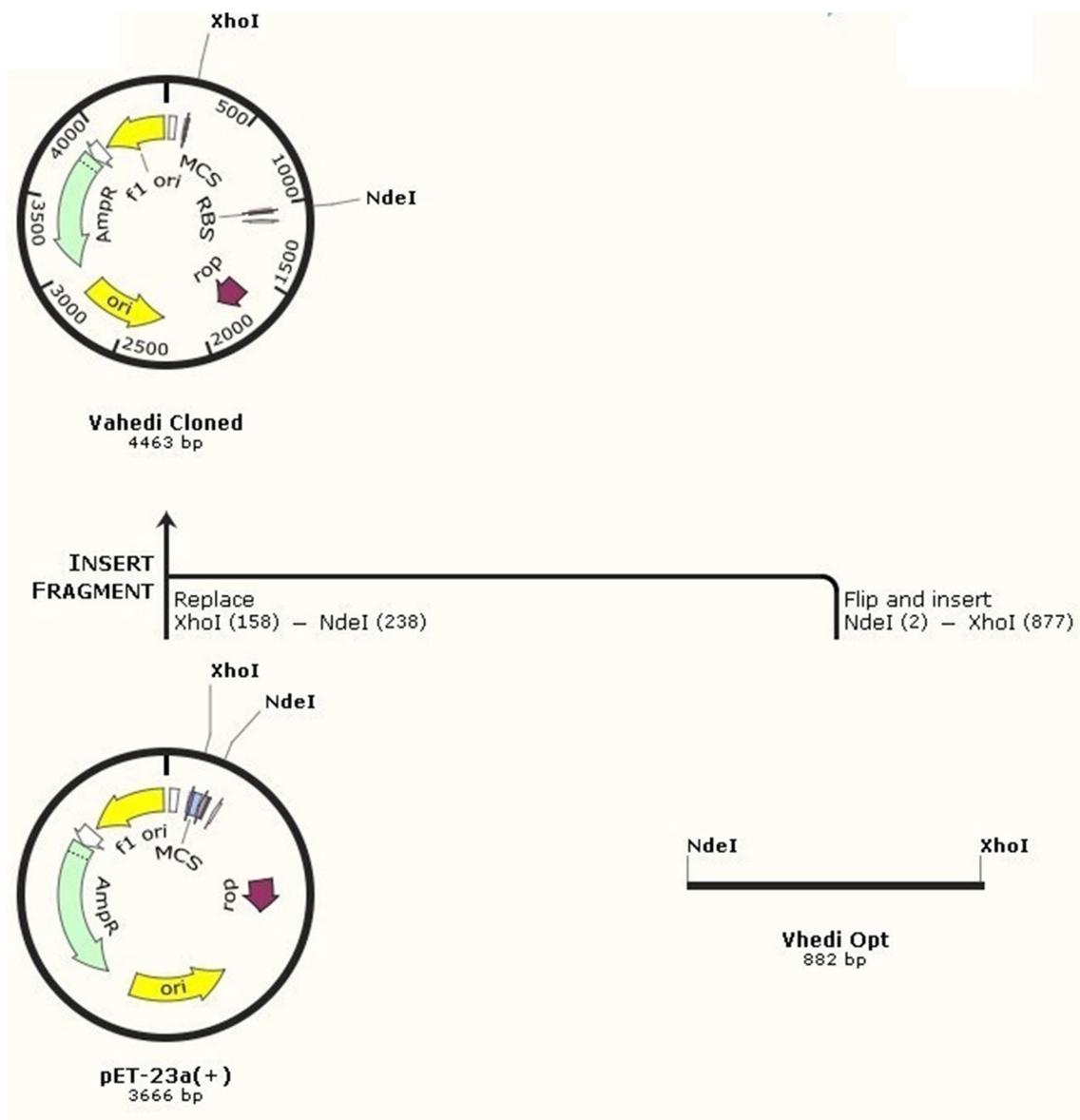


Fig. 8 Schematic representation of cloning process. The gene sequence of multi-epitope protein with the name of Vahedi opt, and the length of 882 nucleotides shown in the designed structure. The

Vahedi opt gene sequence in the Vahedi cloned structure is distinct between the *NdeI* and *XhoI* cleavage sites

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10989-022-10413-1>.

Acknowledgements We are grateful to all members of Department of Medical biotechnology at Shiraz University of Medical Sciences for technical assistance on this work.

Funding This manuscript was extracted from the PhD thesis of Farzaneh Vahedi and was supported by Shiraz University of Medical Sciences, Shiraz, Iran.

Data Availability All data generated or analysed during this study are included in this published article, and its supplementary information files.

Declarations

Conflict of interest All the authors declare that they have none conflict of interest.

Research Involving Human and/or Animal Participants This article does not contain any studies with human participants or animals performed by any of the authors.

References

- Adams AL, Eberle K, Colón JR, Courville E, Xin H (2021) Synthetic conjugate peptide Fba-Met6 (MP12) induces complement-mediated resistance against disseminated *Candida albicans*. *Vaccine* 39:4099–4107
- Antinori S, Milazzo L, Sollima S, Galli M, Corbellino M (2016) Candidemia and invasive candidiasis in adults: a narrative review. *Eur J Intern Med* 34:21–28
- Araújo D, Henriques M, Silva S (2017) Portrait of *Candida* species biofilm regulatory network genes. *Trends Microbiol* 25:62–75
- Bansal AK (2005) Bioinformatics in microbial biotechnology—a mini review. *Microb Cell Fact* 4:1–11
- Bhakt P, Battu A, Kaur R (2018) *Candida*: friend and foe of humans. In: Singh PP (ed) *Infectious diseases and your health*. Springer, Singapore, pp 403–418
- Brown SE et al (2019) The vaginal microbiota and behavioral factors associated with genital *Candida albicans* detection in reproductive-age women. *Sex Transm Dis* 46:753
- Buonfrate D, Formenti F, Perandin F, Bisoffi Z (2015) Novel approaches to the diagnosis of *Strongyloides stercoralis* infection. *Clin Microbiol Infect* 21:543–552
- Calandra T, Roberts JA, Antonelli M, Bassetti M, Vincent J-L (2016) Diagnosis and management of invasive candidiasis in the ICU: an updated approach to an old enemy. *Crit Care* 20:1–6
- Chatterjee R et al (2021) Development of a conserved chimeric vaccine for induction of strong immune response against *Staphylococcus aureus* using immunoinformatics approaches. *Vaccines* 9:1038
- Chen X, Zaro JL, Shen W-C (2013) Fusion protein linkers: property, design and functionality. *Adv Drug Deliv Rev* 65:1357–1369
- Chen H, Zhou X, Ren B, Cheng L (2020) The regulation of hyphae growth in *Candida albicans*. *Virulence* 11:337–348
- Colovos C, Yeates TO (1993) Verification of protein structures: patterns of nonbonded atomic interactions. *Protein Sci* 2:1511–1519
- Dey J, Mahapatra SR, Singh P, Patro S, Kushwaha GS, Misra N, Suar M (2021) B and T cell epitope-based peptides predicted from clumping factor protein of *Staphylococcus aureus* as vaccine targets. *Microb Pathog* 160:105171
- Dey J, Mahapatra SR, Lata S, Patro S, Misra N, Suar M (2022) Exploring *Klebsiella pneumoniae* capsule polysaccharide proteins to design multi-epitope subunit vaccine to fight against pneumonia. *Expert Rev Vaccines*. <https://doi.org/10.1080/14760584.2022.2021882>
- Díez A, Carrano G, Bregón-Villaloz M, Cuétara M-S, García-Ruiz JC, Fernández-de-Larriñoa I, Moragues M-D (2021) Biomarkers for the diagnosis of invasive candidiasis in immunocompetent and immunocompromised patients. *Diagn Microbiol Infect Dis* 101:115509
- Erwig LP, Gow NA (2016) Interactions of fungal pathogens with phagocytes. *Nat Rev Microbiol* 14:163–176
- Gasteiger E, Hoogland C, Gattiker A, Wilkins MR, Appel RD, Bairoch A (2005) Protein identification and analysis tools on the ExPASy server. In: Walker JM (ed) *The proteomics protocols handbook*. Humana Press, Totowa, pp 571–607
- Gharaghani M et al (2018) Identification of *Candida* species isolated from vulvovaginal candidiasis patients by polymerase chain reaction-restriction fragment length polymorphism (PCR-RFLP) in Yasuj southwestern Iran. *Jundishapur J Microbiology* 11(8):e65359
- He Z, Chen J, Li W, Cheng Y, Zhang H-p, Zhang L-n, Hou T-w (2015) Serological response and diagnostic value of recombinant *Candida* cell wall protein enolase, phosphoglycerate kinase, and β -glucosidase. *Front Microbiol* 6:920
- Jespersen MC, Peters B, Nielsen M, Marcatili P (2017) BepiPred-2.0: improving sequence-based B-cell epitope prediction using conformational epitopes. *Nucleic Acids Res* 45:W24–W29
- Khatami SH, Taheri-Anganeh M, Arianfar F, Savardashtaki A, Sarkari B, Ghasemi Y, Mostafavi-Pour Z (2020a) Analyzing signal peptides for secretory production of recombinant diagnostic antigen B8/1 from *Echinococcus granulosus*: an in silico approach. *Mol Biol Res Commun* 9:1
- Khatami SH, Taheri-Anganeh M, Movahedpour A, Savardashtaki A, Ramezani A, Sarkari B, Mostafavi-Pour Z (2020b) Serodiagnosis of human cystic echinococcosis based on recombinant antigens B8/1 and B8/2 of *Echinococcus granulosus*. *J Immunoassay Immunochem* 41:1010–1020
- Ko J, Park H, Seok C (2012) GalaxyTBM: template-based modeling by building a reliable core and refining unreliable local regions. *BMC Bioinform* 13:1–8
- Lain A, Elguezal N, Brena S, García-Ruiz JC, Del Palacio A, Moragues MD, Pontón J (2007a) Diagnosis of invasive candidiasis by enzyme-linked immunosorbent assay using the N-terminal fragment of *Candida albicans* hyphal wall protein 1. *BMC Microbiol* 7:1–11
- Lain A, Moragues MD, Ruiz JCG, Mendoza J, Camacho A, Del Palacio A, Pontón J (2007b) Evaluation of a novel enzyme-linked immunosorbent assay to detect immunoglobulin G antibody to enolase for serodiagnosis of invasive candidiasis. *Clin Vaccine Immunol* 14:318–319
- Lain A, Elguezal N, Amutio E, Fernández de Larriñoa I, Moragues MD, Pontón J (2008) Use of recombinant antigens for the diagnosis of invasive candidiasis. *Clin Dev Immunol*. <https://doi.org/10.1155/2008/721950>
- Lee M, Bang K, Kwon H, Cho S (2013) Enhanced antibacterial activity of an attacin-coleoptericin hybrid protein fused with a helical linker. *Mol Biol Rep* 40:3953–3960
- León C, Ostrosky-Zeichner L, Schuster M (2014) What's new in the clinical and diagnostic management of invasive candidiasis in critically ill patients. *Intensive Care Med* 40:808–819
- Levenhagen MA, Costa-Cruz JM (2014) Update on immunologic and molecular diagnosis of human strongyloidiasis. *Acta Trop* 135:33–43
- Lewis M, Williams D (2017) Diagnosis and management of oral candidosis. *Br Dent J* 223:675–681
- Li F et al (2007) Eap1p, an adhesin that mediates *Candida albicans* biofilm formation in vitro and in vivo. *Eukaryot Cell* 6:931–939
- Li G et al (2016) Construction of a linker library with widely controllable flexibility for fusion protein design. *Appl Microbiol Biotechnol* 100:215–225
- Lombardi L, Zoppo M, Rizzato C, Bottai D, Hernandez AG, Hoyer LL, Tavanti A (2019) Characterization of the *Candida orthopsilosis* agglutinin-like sequence (ALS) genes. *PLoS ONE* 14:e0215912
- Lopéz-Blanco JR, Garzón JI, Chacón P (2011) iMod: multipurpose normal mode analysis in internal coordinates. *Bioinformatics* 27:2843–2850
- Lovell SC et al (2003) Structure validation by α geometry: ϕ , ψ and ω deviation. *Proteins* 50:437–450
- Ma C-f, Li F-q, Shi L-n, Hu Y-a, Wang Y, Huang M, Kong Q-q (2013) Surveillance study of species distribution, antifungal susceptibility and mortality of nosocomial candidemia in a tertiary care hospital in China. *BMC Infect Dis* 13:1–9
- Mahapatra SR, Sahoo S, Dehury B, Raina V, Patro S, Misra N, Suar M (2020) Designing an efficient multi-epitope vaccine displaying interactions with diverse HLA molecules for an efficient humoral and cellular immune response to prevent COVID-19 infection. *Expert Rev Vaccines* 19:871–885
- Mahapatra SR et al (2021) Immunoinformatic approach employing modeling and simulation to design a novel vaccine construct targeting MDR efflux pumps to confer wide protection against

- typhoidal *Salmonella serovars*. J Biomol Struct Dyn. <https://doi.org/10.1080/07391102.2021.1964600>
- Mayer FL, Wilson D, Hube B (2013) *Candida albicans* pathogenicity mechanisms. Virulence 4:119–128
- Nobile CJ et al (2012) A recently evolved transcriptional network controls biofilm development in *Candida albicans*. Cell 148:126–138
- Oliver JC, Ferreira CBRJ, Silva NC, Dias ALT (2019) *Candida* spp. and phagocytosis: multiples evasion mechanisms. Antonie van Leeuwenhoek 112(10):1409–23
- Orsi CF et al (2014) Impact of *Candida albicans* hyphal wall protein 1 (HWP1) genotype on biofilm production and fungal susceptibility to microglial cells. Microb Pathog 69:20–27
- Poissy J et al (2020) Risk factors for candidemia: a prospective matched case-control study. Crit Care 24:1–11
- Ramalho-Oliveira R, Oliveira-Vieira B, Viola JP (2019) IRF2BP2: a new player in the regulation of cell homeostasis. J Leukoc Biol 106:717–723
- Rice P, Longden I, Bleasby A (2000) EMBOSS: the European molecular biology open software suite. Trends Genet 16:276–277
- Richardson JP et al (2018) Processing of *Candida albicans* Ece1p is critical for candidalysin maturation and fungal virulence. mBio. <https://doi.org/10.1128/mBio.02178-17>
- Rodríguez-Cerdeira C et al (2020) Pathogenesis and clinical relevance of *Candida* biofilms in vulvovaginal candidiasis. Front Microbiol 11:2884
- Sabetian S, Nezafat N, Dorosti H, Zarei M, Ghasemi Y (2019) Exploring dengue proteome to design an effective epitope-based vaccine against dengue virus. J Biomol Struct Dyn 37:2546–2563
- Sachivkina N, Lenchenko E, Blumenkrants D, Ibragimova A, Bazarkina O (2020) Effects of farnesol and lyticase on the formation of *Candida albicans* biofilm. Vet World 13:1030
- Sáez-Rosón A, Sevilla M-J, Moragues M-D (2014) Identification of superficial *Candida albicans* germ tube antigens in a rabbit model of disseminated candidiasis. A proteomic approach. Int Microbiol 17:21–29
- Saha S, Raghava GPS (2004) BcePred: prediction of continuous B-cell epitopes in antigenic sequences using physico-chemical properties. International conference on artificial immune systems. Springer, Berlin, pp 197–204
- Saha S, Raghava GPS (2006) Prediction of continuous B-cell epitopes in an antigen using recurrent neural network. Proteins 65:40–48
- Sefidi FJ, Keyvani H, Monavari SH, Alavian SM, Fakhim S, Bokharai-Salim F (2013) Distribution of hepatitis C virus genotypes in Iranian chronic infected patients. Hepat Mon 13:e7991
- Shahbazi M, Haghkhal M, Rahbar MR, Nezafat N, Ghasemi Y (2016) In silico sub-unit hexavalent peptide vaccine against an *Staphylococcus aureus* biofilm-related infection. Int J Pept Res Ther 22:101–117
- Sherry L et al (2014) Biofilms formed by *Candida albicans* bloodstream isolates display phenotypic and transcriptional heterogeneity that are associated with resistance and pathogenicity. BMC Microbiol 14:1–14
- Shukla M, Rohatgi S (2020) Vaccination with secreted aspartyl proteinase 2 protein from candida parapsilosis can enhance survival of mice during *C. tropicalis*-mediated systemic candidiasis. Infect Immun 88:e00312–00320
- Singh DK et al (2019) Functional characterization of secreted aspartyl proteases in *Candida parapsilosis*. Msphere 4:e00484–00419
- Souza MN, Ortiz SO, Mello MM, Oliveira FdM, Severo LC, Goebel CS (2015) Comparison between four usual methods of identification of *Candida* species. Rev Inst Med Trop São Paulo 57:281–287
- Tehrani SS et al (2020) Suitable signal peptides for secretory production of recombinant granulocyte colony stimulating factor in *Escherichia coli*. Recent Pat Biotechnol 14:269–282
- Thomson DD, Wehmeier S, Byfield FJ, Janmey PA, Caballero-Lima D, Crossley A, Brand AC (2015) Contact-induced apical asymmetry drives the thigmotropic responses of *Candida albicans* hyphae. Cell Microbiol 17:342–354
- Tsui C, Kong EF, Jabra-Rizk MA (2016) Pathogenesis of *Candida albicans* biofilm. Pathog Dis. <https://doi.org/10.1093/femsdp/ftw018>
- Uppuluri P et al (2018) The Hyr1 protein from the fungus *Candida albicans* is a cross kingdom immunotherapeutic target for Acinetobacter bacterial infection. PLoS Pathog 14:e1007056
- Wang X-j, Sui X, Yan L, Wang Y, Cao Y-b, Jiang Y-y (2015) Vaccines in the treatment of invasive candidiasis. Virulence 6:309–315
- Wiederstein M, Sippl MJ (2007) ProSA-web: interactive web service for the recognition of errors in three-dimensional structures of proteins. Nucleic Acids Res 35:W407–W410
- Zarei M, Nezafat N, Rahbar MR, Negahdaripour M, Sabetian S, Morowvat MH, Ghasemi Y (2019) Decreasing the immunogenicity of arginine deiminase enzyme via structure-based computational analysis. J Biomol Struct Dyn 37:523–536

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.